Individual Differences in Perception of Foreign-Accented Speech: The Role of Executive Function

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in Speech and Hearing Science in the undergraduate colleges of The Ohio State University

by

Angela S. Hill

The Ohio State University

April 2015

Project Advisor: Dr. Rachael Frush Holt, Ph.D., Department of Speech and Hearing Science

Table of Contents

Abstract

Large individual differences in spoken word recognition have been observed in children with hearing loss and with normal hearing. Executive functions account for some of this variability (Beer, Kronenberger, & Pisoni, 2011; Lalonde & Holt, 2014). Individual performance differences are also evident when listeners are presented with nonnative speech. This study examined the influence of executive function on children's nonnative speech perception. Eighty-four 5- to 7-year-old monolingual English-speaking children were presented with 60 English sentences produced by either a native English or Mandarin-accented talker (Van Engen et al., 2010) embedded in multi-talker babble at +8 dB signal-to-noise ratio. For 30 sentences, the final (target) word was highly predictive from sentence context and for the other 30 it was not; the same final words appeared in both predictability conditions. Sentence context facilitated target word recognition similarly for native and nonnative talkers. Children with better inhibition, as assessed by the Behavior Rating Inventory of Executive Function (Gioia, Isquith, Guy, & Kenworthy, 2000), had better perception of nonnative speech than those with poor inhibition. Stronger executive functions appear to support word recognition under adverse listening conditions, including both those stemming from the listener (e.g., hearing loss) or, as shown here, the talker (e.g., nonnative speech).

Chapter 1

Introduction

Foreign-Accented Speech Perception

Speech perception refers to how spoken language is processed and interpreted by a listener (Poeppel & Monahan, 2008). To be comprehended, the auditory signal must be transmitted into a mental representation that the person can perceive (Poeppel & Monahan, 2008). To do so, one must cope with a constant in speech: variability (McLennan & Luce, 2005). There are many sources of variability in speech, such as speaking rate and speaking style. This study will examine one of them - foreign accent.

Variability in foreign-accented speech creates challenges to distinguishing fine phonetic details (e.g., idiolect differences and positional effects), as well as segmental and suprasegmental characteristics (e.g., substitutions, deletions, additions, stress, rhythm, and intonation differences) compared to normal variability in speech (Bent & Holt, 2013). Although a foreign-accented speaker has more variability relative to a native speaker, the phonetic details and segmental and suprasegmental domains are still organized and structured in the correct way (Bent & Holt, 2013). The phonetic relevance hypothesis proposes that not all sources of variability affect speech perception equally; specifically, acoustic properties affect speech perception (Sommers & Barcroft, 2006). When a listener hears foreign-accented speech, the listener has to be able to change how the acoustic information is being processed (Sumner, 2011). This requires greater processing time for nonnative speech than native speech (Adank, Evans, Stuart-Smith, & Scott, 2009; Munro & Derwing, 1995). Consequently, the challenge for listeners is seen in both the variability and instability of foreign-accented speech (Van Engen et al., 2010). The foreign-accented talker's characteristics and speaking rate increase the variability of the sound structure

(Bent & Holt, 2013), causing decreases in spoken word recognition performance relative to native speech in both adults and children (Bent, 2014; Bent & Holt, 2013; Bradlow & Pisoni, 1999; Munro & Derwing, 1995), as well as increases in processing time (Adank et al., 2009; Munro & Derwing, 1995).

Context Effects

Although children have more difficulty perceiving speech in noise overall than adults, they show similar benefits from contextual cues – lexical, semantic, and syntactic – compared to adults (Fallon, Trehub, & Schneider, 2002). Words presented in a high-context (predictable) sentence facilitate more accurate spoken word recognition than those presented in a low-context sentence in adults (Bradlow & Alexander, 2007). However, context is not as useful for nonnative adult listeners as it is for native adult listeners. In particular, nonnative adult listeners required a clearer acoustic signal to capitalize on context effects in noise than did native listeners (Bradlow & Alexander, 2007). Whether young children are able to use context in their recognition of foreign-accented speech is relatively unexplored.

Executive Function

Executive function (EF) is an umbrella term that encompasses many of the cognitive processes that are used to regulate an individual's actions and behaviors (Anderson, 2002; Fuster, 1997; Miller & Cohen, 2001; Shallice & Burgess, 1996). Examples of cognitive functions that fall under the purview of EF include inhibition, shifting, emotional control, initiative, working memory, and planning (Gioia et al., 2000). Some studies refer to executive function as a processing ability that helps an individual resolve conflicting information. Working memory and inhibition are classic examples of this. Working memory is the ability to hold information in one's mind while working on completing an unrelated task. Inhibition is the ability to resist impulses, stop one's own behavior at the appropriate time, and focus on relevant stimuli in the presence of irrelevant ones (Davidson, Amso, Anderson, & Diamond, 2006; Miyake et al., 2000; Zelazo & Miller, 2002). Both working memory (Lalonde & Holt, 2014) and inhibition (Janse & Adank, 2012) are key components of EF that have been shown to account for some of the variance in speech recognition.

Executive Function in Children with Hearing Impairment

Beer, Kronenberger, and Pisoni (2011) evaluated areas of executive function vulnerability in children with cochlear implants (CIs) by using the Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000) and explored whether they were associated with the children's speech and language outcomes. They used the Clinical Evaluation of Language Fundamentals (CELF-4; Semel, Wiig, Secord, 2003) to assess general language ability, the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007) to assess vocabulary, the Lexical Neighborhood Test (LNT; Kirk, Pisoni, & Osberger, 1995) to assess open-set word recognition, and the Hearing in Noise Test for Children (HINT-C; Nilsson, Soli, & Sullivan, 1994) to assess open-set sentence recognition in quiet and noise. They found that children with CIs had difficulty on specific domains of EF: inhibition and working memory. Importantly, spoken word recognition in noise and language could be predicted by children's working memory.

A different study by Kronenberger, Pisoni, Henning, and Colson (2013) investigated whether there were differences in executive function between normal-hearing individuals and long-term CI users. A battery of tests was used to measure executive function (e.g., Retrieval Fluency subtest of the Woodcock–Johnson Tests of Cognitive Abilities, Third Edition [WJ-III; Woodcock, McGrew, & Mather, 2001]; Digit Span subtest of the Wechsler Intelligence Scale for

Children, Third Edition [WISC-III; Wechsler, 1991]; Test of Variables of Attention [TOVA; Leark, Dupuy, Greenberg, Corman, & Kindschi, 1996]). They found that a significant number of long-term CI users may not be able to "catch up" with their normal-hearing peers on certain areas of EF: working memory, fluency (another EF domain not used in the BRIEF), and inhibition were all found to be one standard deviation below the normative mean, suggesting that long-term cochlear implant users may still have significant cognitive delays even after having a CI for at least 7 years.

Development of Executive Function

EF's cognitive and self-regulatory processes are related to the development of the prefrontal cortex (Miller & Cohen, 2001; Stuss & Knight, 2002). During childhood, the prefrontal cortex and EF are still developing, and continue to do so through early adulthood (Blakemore & Choudhury, 2006). EF, which is controlled by the prefrontal cortex, impacts the perception of speech (Miller & Cohen, 2001; Stuss & Knight, 2002). Difficulties with inhibition, working memory, and auditory attention have been linked to poor spoken language outcomes in children with CIs (e.g., Pisoni, Conway, Kronenberger, Henning, & Anaya, 2010). Recently, Lalonde and Holt (2014) showed that working memory was related to speech discrimination in normal-hearing toddlers.

Studies have intensely researched executive function due to its importance in child and adolescent brain development, but less so in its specific influence on perception of foreign-accented speech. The present study will draw from these findings and will contribute new knowledge to the field to help inform future research in this area.

Chapter 2

Hypothesis

The purpose of this investigation was to examine children's perception of foreign-accented speech and how executive function influences individual differences in its perception. We predicted that children with better working memory and inhibition will be able to identify and repeat back more words in sentences correctly from nonnative speakers than children with the opposite characteristics. A secondary goal was to evaluate children's use of context in nonnative speech, and whether the size of the benefit differs between native and nonnative speech. We predicted that recognition facilitated by context will be better than that which is not.

Chapter 3

Method

Participants

Data from eighty-four 5- to 7-year-old children were used in this study. The children were all monolingual, American English speaking, with a parent-reported history of typical speech, hearing, and language development. Twelve additional participants were excluded due to exposure to Mandarin-accented English (the foreign accent used in this study) or other foreign-accented speech (n=8), parent-reported history of speech problems (n=2), or a parent not filling out a questionnaire (n=2). The mean age of the 5-year-olds was 5.45 (SD = 0.32), 6-year-olds was 6.45 (SD = 0.29), and 7-year-olds was 7.47 (SD = 0.29) years. There were: 31 5-year-olds, of which 15 were female and 16 were male; 36 6-year-olds, including 21 females and 15 males; and 29 7-year-olds, of which 15 were female and 14 were male. The Institutional Review Board at The Ohio State University approved the use of human subjects for this project.

Equipment

Custom Python software running on a Dell Optiplex 790 computer controlled the experiment. Stimuli were presented via audio-technica headphones (model number ETH-770COM) binaurally.

Stimuli

A list of 60 pre-recorded English sentences of speakers of different language backgrounds from The Wildcat Corpus of native and foreign-accented English (Van Engen et al., 2010) embedded in 8-talker-babble of +8 dB SNR were used as the stimuli. Two talkers were selected: a native American English male and a native Mandarin Chinese-accented male. On half of the sentences, the final word could be predicted from the context of the sentences (high context); on the other half of the sentences the final word could not be predicted from the sentence context (low context). An example of a low context sentence is, “We pointed to the animals.” An example of a high context sentence is, “Elephants are big animals.” Each final word appeared twice in the set of sentences (once in low- and once in high-predictability).

Materials

Behavior Rating Inventory of Executive Function (BRIEF)

The BRIEF (Gioia et al., 2000) is a psychometrically rigorous parental report of EF. It is a validated 86-item questionnaire in which a single caregiver indicates if a list of behaviors were a problem or not for their child in the last six months. The BRIEF includes eight scales: inhibition, shifting, emotional control, initiation, working memory, planning/organization, organization of materials, and monitoring.

Language Background Questionnaire

The language background questionnaire required the parents to give their child’s age, gender, ethnicity, past and current places of residency, if their child had any speech or hearing problems, and when their child began speaking. If the child knew more than one language, the parents were asked to fill in what percentage of the time each language was used/spoken. The parents also indicated the amount of exposure their child had to other languages or accents on a scale from 1-5, with 1 being no exposure to 5 being daily exposure.

Procedure

Children were randomly assigned to either the native or nonnative speaker. Before testing, children were presented with four practice sentences (two high-predictability and two low-predictability, one of each spoken by each of the two speakers). Then they were tested on all 60 sentences presented in random order under headphones at a comfortable listening level and were asked to repeat each sentence, guessing if necessary. Only the final word was scored. Children received non-contingent praise after each item. The +8 dB SNR babble preceded and followed the sentence by 500ms. While the child completed the speech recognition task, parents filled out the language background questionnaire and the BRIEF. The total testing time was approximately 15 minutes. Participants were not paid. The research was conducted at the Center of Science and Industry (COSI) in Columbus, Ohio.

Chapter 4

Results

Speech Perception

Mean percent correct and +1 standard deviation for final words in sentences as a function of speaker accent are displayed in Figure 1. A 3-way mixed effects ANOVA (within subjects factor: context; between subjects factors: age and accent) revealed that the children identified words spoken by the native speaker more accurately than those by the nonnative speaker, $F\ (1,78) = 95.847$, $p < 0.0001$. Figure 2 displays mean percent correct and +1 standard deviation for final words in sentences as a function of context. Words in high-predictability contexts were identified more accurately than those in low-predictability contexts, $F\ (1,78) = 53.793$, $p < 0.0001$. Average performance as a function of age is displayed in Figure 3. As expected, older children achieved better accuracy than younger children, $F\ (2,78) = 6.841$, $p = 0.002$. Post-hoc testing with a conservative Bonferroni correction revealed that 7-year-olds had significantly better recognition accuracy than 5-year-olds ($p = .001$) There was a significant interaction between context and age, $F\ (2, 78) = 11.032$, $p < .0001$, which is displayed in Figure 4: the performance benefit obtained by having context cues for words presented in high-predictability contexts as opposed to low-predictability contexts was larger for older children than younger children.

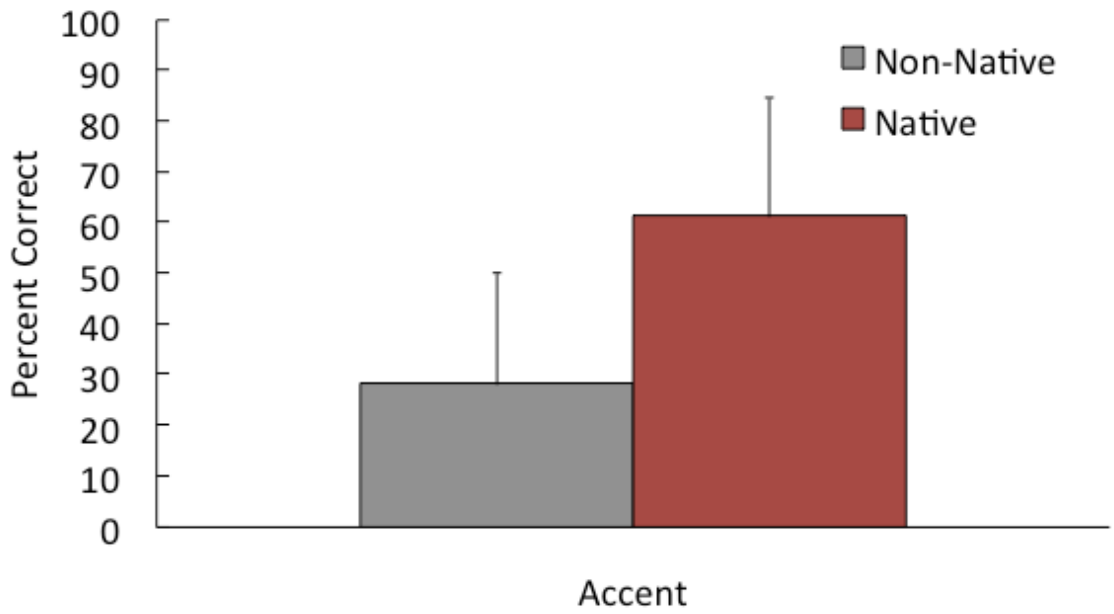

Figure 1. Mean percent final word correct (+ 1 SD) as a function of accent.

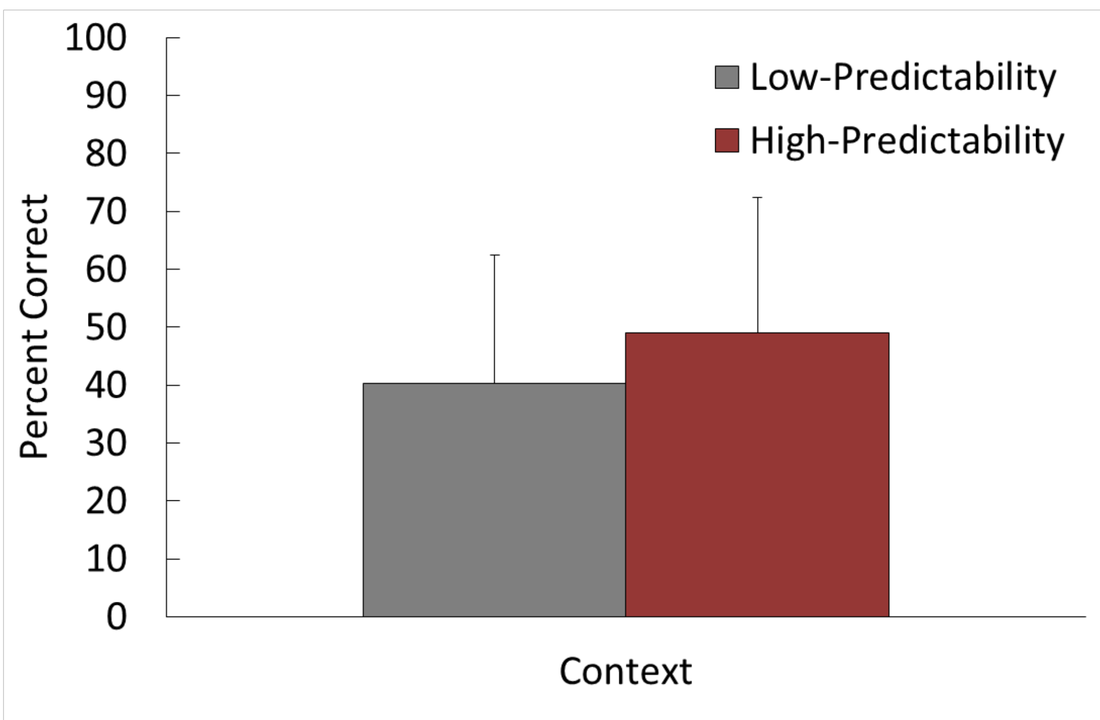

Figure 2. Mean percent correct (+1 SD) as a function of context.

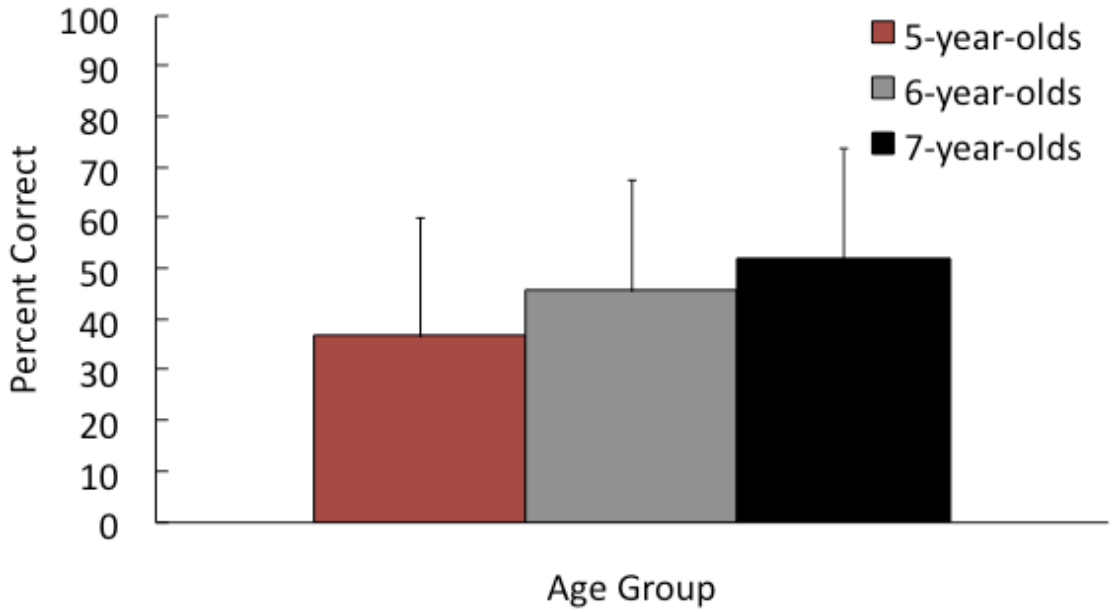

Figure 3. Mean percent correct (+1 SD) as a function of age.

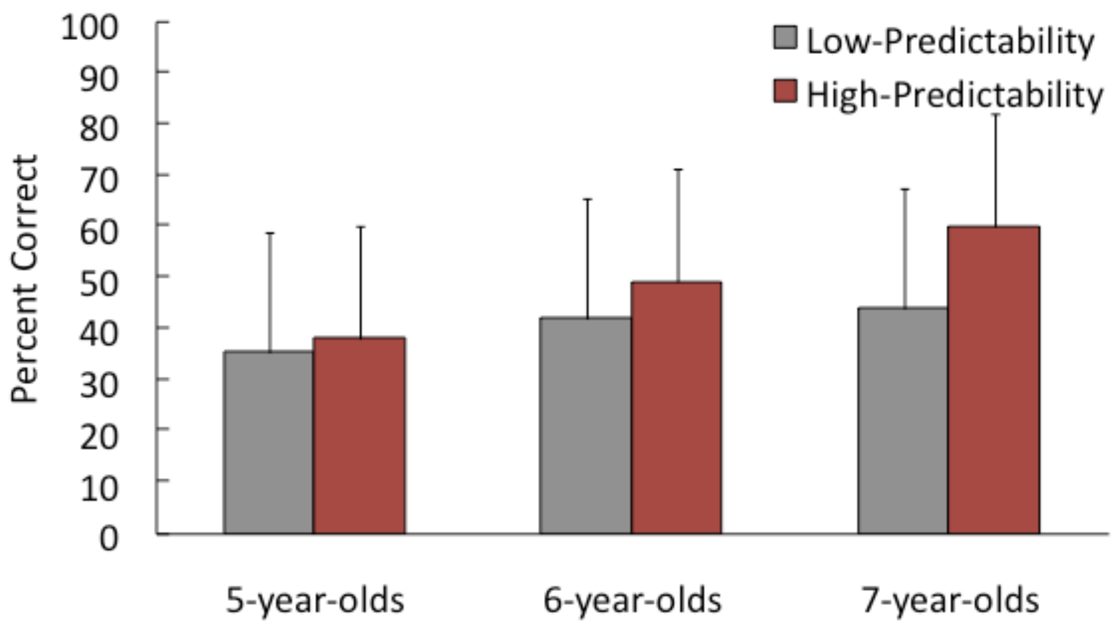

Figure 4. Mean percent correct (+1 SD) as a function of context and age.

Executive Function

Mean *T*-scores (+/- 1 SD) on the BRIEF are shown in Figure 5. A *T*-score of 50 on any subscale is average and anything above a 65 is considered to be clinically elevated, while anything below a 35 is considered clinically depressed. The average scores were within the normal range. To examine whether EF accounted for some of the variability in perception of foreign-accented speech, the speech perception and EF data were analyzed using a stepwise linear regression. Spoken word recognition data were collapsed across high- and low-predictability conditions for these analyses. Data for listeners assigned to the native speaker and those assigned to the nonnative speaker were analyzed separately. The regression model was only significant for the data from the nonnative speaker. These results are displayed in Table 1. The only subscale of the BRIEF that was significant in the regression model was Inhibit ($p < .05$): children whose parents reported that they had more problems with inhibitory control had poorer word-in-sentence recognition scores than those whose parents reported fewer problems in this EF domain. In fact, parent's ratings of children's inhibitory control accounted for approximately 10% of the variability in children's nonnative spoken word recognition scores. There were no EF domains that accounted for any significant amount of variability in native-speaker spoken word recognition.

Table 1. Stepwise linear regression results for the nonnative speaker, speech recognition data averaged across context levels.

	B	SE B	Beta
Constant	54.435	12.473	
BRIEF Inhibit	-.495	.232	-.320*

Notes: $R^2 = .102$, **p* <. 05.

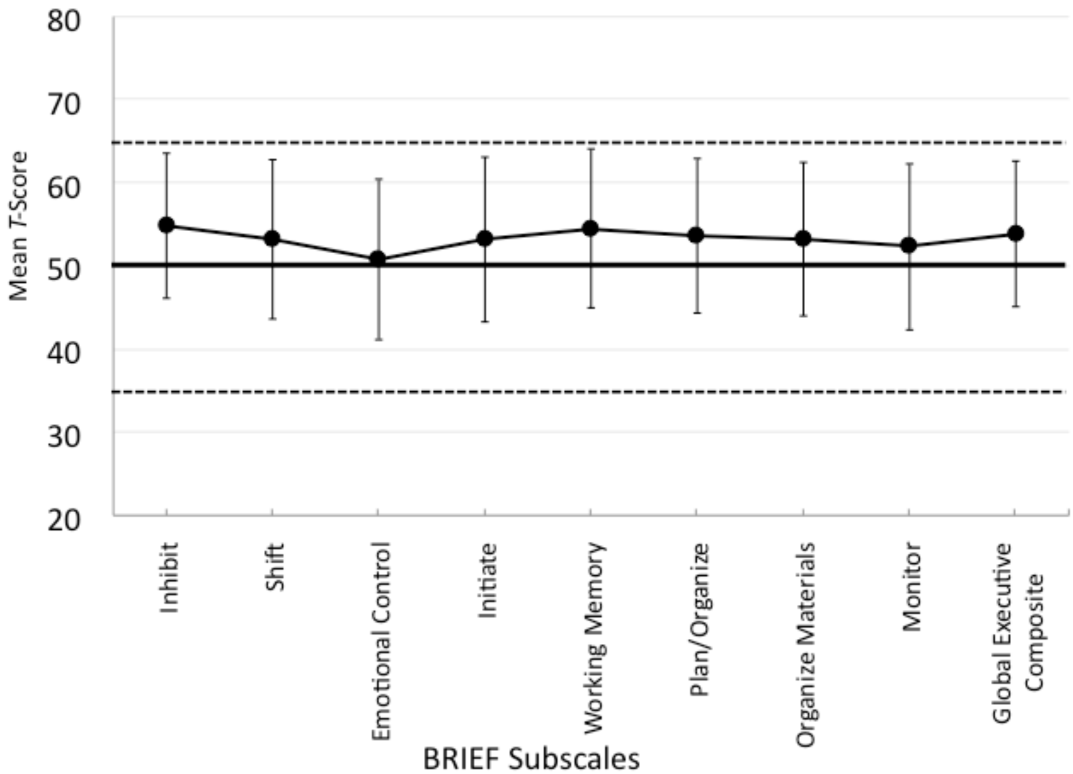

Figure 5. Mean *T*-scores (+/- 1SD) on the BRIEF. Higher scores indicate more parent-reported difficulties.

Chapter 5

Discussion and Conclusion

Children were able to capitalize on sentence context when presented with foreign-accented speech in noise, extending previous findings for native-accented speech in noise (Fallon et al., 2002). Greater first language experience might facilitate the use of contextual cues: context effects are more beneficial to older children than younger children. Similarly, adult second-language learners are less able to take advantage of context than native listeners, because they have less experience with the target language (Bradlow & Alexander, 2007).

The results from the regression analysis support and extend those reported recently from older adults: adaptation to an artificial novel accent could be partially predicted by their selective attention (measured by the flanker task; Janse & Adank, 2012). The task required participants to inhibit incorrect responses in incongruent conditions. Together, these studies suggest that inhibition appears to contribute to individual differences in both children's and adults' perception of unfamiliar accents. Children and adults who have better inhibitory control appear to be better at identifying words produced by speakers of an unfamiliar accent.

It appears that inhibitory control allows the listener to do better at perceiving the foreign-accented speaker in noise because of their ability to ignore the background noise and the foreign accent-induced differences in phonetic detail and segmental and suprasegmental domains. Children who have the ability to focus on the relevant auditory cues are better able to process the acoustic information coming from the foreign-accented speaker more effectively than children who are less able to inhibit cues that do not help with the task. Without strong inhibitory skills, children are poorer at focusing on the relevant talker's cues and have a harder time disregarding the unimportant, irrelevant information.

Future studies should investigate whether these results apply to speakers of other foreign accents, as well as speakers of different regional dialects/accents. Additionally, the gender of the talker should be taken into account. Only a male speaker was used in the current study; future studies should assess if female voices have any different outcomes.

As for executive function, future studies could use a performance task to measure the different subscales. In the current study, the BRIEF, a parent-report questionnaire, was used to assess the child's EF. Performance measures would be another way to quantify executive function and would shed additional light on these findings.

Acknowledgements

We are grateful to Charles Brandt for programming, the families who participated in this research, COSI and the BLN for their facilities, the Social and Behavioral Sciences Undergraduate Research grant from OSU, OSU’s Undergraduate Research Scholar award, and OSU’s Arts and Sciences College Undergraduate Research Scholarship for the financial support for this project. I would also like to thank my advisor, Dr. Holt, for her tremendous support, time, and dedication into this project.

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(2), 520–9.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *8*(2), 71–82. http://doi.org/10.1076/chin.8.2.71.8724

Beer, J., Kronenberger, W. G., & Pisoni, D. B. (2011). Executive function in everyday life: implications for young cochlear implant users. *Cochlear Implants International*, *12*(S1), S89–S91. http://doi.org/10.1179/146701011X13001035752570

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language*, *41*(6), 1334–1355.

Bent, T., & Holt, R. F. (2013). The influence of talker and foreign-accent variability on spoken word identification. *The Journal of the Acoustical Society of America*, *133*(3), 1677–1686. http://doi.org/10.1121/1.4776212

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *JCPP Journal of Child Psychology and Psychiatry*, *47*(3-4), 296–312.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339–2349.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, *106*(4), 2074–85.

Davidson, M., Amso, D., Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia Neuropsychologia*, *44*(11), 2037–2078.

Dunn, L., & Dunn, L. (2007). Peabody Picture Vocabulary Test- Fourth Edition. Bloomington, MN: Pearson Assessments.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2002). Children's use of semantic cues in degraded listening environments. *The Journal of the Acoustical Society of America*, *111*(5), 2242–2249. http://doi.org/10.1121/1.1466873

Fuster, J. M. (1997). *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe*. Philadelphia: Lippincott-Raven.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *6*(3), 235–238. http://doi.org/10.1076/chin.6.3.235.3152

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology (2006)*, *65*(8), 1563–85.

Kirk, K. I., Pisoni, D. B., & Osberger, M. J. (1995). Lexical Effects on Spoken Word Recognition by Pediatric Cochlear Implant Users. *Ear & Hearing*, *16*(5), 470–481.

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., & Colson, B. G. (2013). Executive functioning skills in long-term users of cochlear implants: a case control study. *Journal of Pediatric Psychology*, *38*(8), 902–14.

Lalonde, K., & Holt, R. F. (2014). Cognitive and linguistic sources of variance in 2-year-olds' speech-sound discrimination: a preliminary investigation. *Journal of Speech, Language, and Hearing Research: JSLHR*, *57*(1), 308–326. http://doi.org/10.1044/1092-4388(2013/12-0227)

Leark, R. A., Dupuy, T. R., Greenberg, L. M., Corman, C. L., & Kindschi, C. L. (1996). *Test of variables of attention professional manual version 7.0*. Los Alimitos, CA: Universal Attention Disorders, Inc.

McLennan, C.T., & Luce, P.A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(2), 306–21.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and noise. *Journal of Acoustical Society of America*, *95*(2), 1085-1098.

Pisoni, D. B., Conway, C. M., Kronenberger, W., Henning, S., & Anaya, E. (2010). Executive function and cognitive control in deaf children with cochlear implants. In M.S. Marschark & P.E. Spencer (Eds.), Oxford handbook of deaf studies, language, and education (2nd ed., Vol. 1, pp. 439-457). New York, NY: Oxford University Press.

Poeppel, D., & Monahan, P. J. (2008). Speech Perception: Cognitive Foundations and Cortical Implementation. *CDIR Current Directions in Psychological Science*, *17*(2), 80–85.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical Evaluation of Language Fundamentals (CELF – 4th Edition). San Antonio, TX: Harcourt Assessment.

Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1405–11.

Sommers, M. S., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *The Journal of the Acoustical Society of America*, *119*(4), 2406–2416.

Stuss, D. T., & Knight, R. T. (2002). *Principles of frontal lobe function*. Oxford; New York: Oxford University Press.

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition*, *119*(1), 131–136. http://doi.org/10.1016/j.cognition.2010.10.018

Van Engen, K., Baese-Berk, M., Baker, R., Choi, A., Kim, M., & Bradlow, A. (2010). The Wildcat Corpus of Native-and Foreign-accented English: Communicative Efficiency across Conversational Dyads with Varying Language Alignment Profiles. *Language and Speech*, *53*(4), 510–540.

Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maerlender, A. (2004). *Wechsler intelligence Scale for Children–Fourth Edition–Integrated*. San Antonio, TX: Harcourt Assessment.

Woodcock, R. W., McGrew, K. S., & Mather, J. (2001). *Woodcock-Johnson-III*. Itasca, IL: Riverside Publishing.

Zelazo, P. D., & Miller, U. (2002). Executive Function in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp.445-469). Oxford, England: Blackwell. doi:10.1002/9780470996652.ch20